S61282 C2

N63 17231
Code 1 39 P

gp note
OSP, LRC

TECHNICAL MEMORANDUM NO. 25

DETECTION OF SIGNALS IN COLORED NOISE

Bernard Harris

May 1, 1962

OTS PRICE

XEROX $ ~~3.60 ph~~

MICROFILM $ ~~1.22 mf~~

NEW YORK UNIVERSITY
COLLEGE OF ENGINEERING
DEPARTMENT OF ELECTRICAL ENGINEERING

University Heights
New York 53, New York

NASA CR-50348

TECHNICAL MEMORANDUM NO. 25

To: Cyrus J. Creveling, Code 561
National Aeronautics and Space Administration
Greenbelt, Maryland

Re: Contract No. NAS 5-408

Subject: Detection of Signals in Colored Noise

SUMMARY

The purpose of this memorandum is twofold: to extend the concepts discussed in Technical Memorandum No. 2 and to lay the groundwork for Technical Memorandum No. 26. Technical Memorandum No. 2 was concerned with the basic concepts of signal design and signal reception when the channel interference is additive white gaussian noise. By white noise we mean that the noise power per unit bandwidth is the same at all frequencies of interest. In this case the key results are:

1. The optimum receiver should use correlation techniques (or the equivalent) in which the received signal is multiplied with each of the possible signals that could have been sent. The product signals are then averaged over the signal duration. That correlation process yielding the greatest output at the end of the signal duration indicates which signal is most likely to have been the one sent.

Equivalent techniques are in some cases the same as synchronous detection, while in others the same as filtering (matched filter).

2. The design of optimum signals, when the interference is white noise, is based on the distance parameter

$$D = \int_0^T [s_0(t) - s_1(t)]^2 \, dt$$

which indicates the distinguishability between two signals $s_0(t)$ and $s_1(t)$ as an equivalent amount of energy per message. For example, if PSK (180°) is used and the sinusoids are of peak amplitude A,

$$D[PSK\ (180^\circ)] = 2A^2\ T$$

If FSK is used

$$D[FSK] = A^2\ T$$

Hence PSK (180°) signals have 3db more distinguishability for the same power and duration as FSK signals.

When the interference is additive gaussian, but not white (i.e. the noise power per unit bandwidth is not the same at all frequencies), correlation techniques are again found to be optimum, but the received signal is correlated not with the possible message signals, but with modified signals. These modified signals are obtained from the characteristics of the message signals and the noise power spectrum. The form of these modified signals is not obvious, although the results obtained for specific cases agree with intuitive notions. The key results are as follows.

Consider that there are two possible signals the transmitter can send, $s_0(t)$ and $s_1(t)$. Let the noise power spectrum be $W(\omega)$ and the corresponding autocorrelation function be $R(\tau)$. Now if $y(t)$ is the received signal, the optimum receiver is based on correlation - i.e. forming

$$\int_0^T y(t)\, f(t)\, dt$$

in which T is the signal duration and $f(t)$ is the solution of the integral equation

$$\int_0^T R(t-\tau)\, f(\tau)\, d\tau = s_0(t) - s_1(t)$$

Basic to the solution of this equation is the corresponding homogeneous integral equation

$$\int_0^T R(t-\tau)\, \varphi(\tau)\, d\tau = \sigma^2\, \varphi(t)$$

As will be discussed in Technical Memorandum No. 26, signals found from sums of those waveforms which satisfy this latter equation have neat properties. One significant property is that bandwidth limitations can be imposed on the selection of signal waveforms by selecting an appropriate noise power spectrum, albeit that the true noise power spectrum is white. In effect the selected noise power spectrum (chosen cup-shape) acts as a weighting function that concentrates the signal power into a narrow range of frequencies.

DISCUSSION

I. Introduction

The decision problem encountered at the receiver of a communication system may be formalized by considering that the receiver is trying to determine which of two possible signals is more likely to have been the transmitted signal. Analysis of this type of binary decision can be extended to cover multi-signal transmissions by simply considering the possibilities two at a time. For this reason we will therefore consider that there are only two possible signals the transmitter can send, $s_0(t)$ and $s_1(t)$.

As discussed in Technical Memorandum No. 2, in a broad sense the receiver is a computer - although in many cases a fairly simple analog computer. From the incoming signal, $y(t)$, the receiver calculates which possible transmitted signal, $s_0(t)$ or $s_1(t)$, more likely would result in $y(t)$ being received. Formally the probability that $s_0(t)$ sent results in $y(t)$ being received, $p[y(t)/s_0(t)]$, and the probability that $s_1(t)$ sent results in $y(t)$ being received, $p[y(t)/s_1(t)]$, are calculated and the likelihood ratio

$$\frac{p[y(t)/s_0(t)]}{p[y(t)/s_1(t)]}$$

is considered. This ratio (or more usually some function of it) is present in the receiver as a voltage level. For completely automatic operation this voltage level is compared with a preset threshold as a means of deciding whether $s_o(t)$ or $s_1(t)$ was sent. Alternatively, the voltage level can be presented to an operator for the final decision. The choice of the threshold level depends on the degree to which each signal is a priori expected (if this is known), the relative costs of making errors, or some other value judgment. For example, as discussed in Technical Memorandum No. 2, a threshold level can be used which corresponds to minimum error probability.

Since the noise is additive the conditional probability of the form $p[y(t)/s(t)]$ is simply the probability that the noise waveform is the difference $[y(t) - s(t)]$. The decision as to whether $s_o(t)$ or $s_1(t)$ was more likely the signal sent is therefore the same as deciding if the noise waveform was more likely $n_o(t) = [y(t) - s_o(t)]$, or $n_1(t) = [y(t) - s_1(t)]$.

To use this fact the possible signals $s_o(t)$ and $s_1(t)$, and the received signal $y(t)$ are sampled at instants of time t_Δ apart such that there are N samples over the waveforms of duration T. The channel is presumed distortionless so that the k^{th} sample of the signal sent causes the k^{th} sample of the received signal. If we let

s_{ok} be the value of $s_o(t)$ at $t = kt_\Delta$

s_{1k} be the value of $s_1(t)$ at $t = kt_\Delta$

y_k be the value of $y(t)$ at $t = kt_\Delta$

then

$$p[y(t)/s_o(t)] = \lim_{t_\Delta \to 0} p[y_1, y_2, \dots y_N / s_{o1}, s_{o2}, \dots s_{oN}]$$

is the same as expecting the interference to have the sampled values

$$\begin{aligned} n_1 &= y_1 - s_{o1} \\ n_2 &= y_2 - s_{o2} \\ &\cdots\cdots \\ n_N &= y_N - s_{oN} \end{aligned}$$

Similarly,

$$p[y(t)/s_1(t)] = \lim_{t_\Delta \to 0} p[y_1, y_2, \dots y_N / s_{11}, s_{12}, \dots s_{1N}]$$

is the same as expecting the interference to have the sampled values

$$\begin{aligned} n_1 &= y_1 - s_{11} \\ n_2 &= y_2 - s_{12} \\ &\cdots\cdots \\ n_N &= y_N - s_{1N} \end{aligned}$$

Comment on rigor: This procedure is satisfactory if the total noise power is finite. If the total noise power is infinite, the "samples" are integrated values of the signals over each duration t_Δ. Alternatively, the noise power density spectrum can be truncated at some high frequency f. Then as $t_\Delta \to 0$, $f \to \infty$.

For the case in which the interference is white gaussian noise the sampled values of the interference are independent and the analysis is greatly simplified. In this case, the joint probability that the sampled values of the noise are $n_1, n_2, \ldots n_k \ldots n_N$ can be written as the product of the individual or marginal probabilities. Thus

$$p(n_1, n_2, \ldots n_k \ldots n_N) = p(n_1)\, p(n_2) \ldots p(n_k) \ldots p(n_N)$$

for which

$$p(n) = \frac{1}{\sqrt{2\pi W^2 f}}\, e^{-n^2/2W^2 f}$$

since p(n) has a Gaussian distribution and the noise power per unit bandwidth is W^2. The remainder of the procedure for determining the optimum receiver is straightforward (See Technical Memorandum No. 2),except for simplicity the log of the likelihood ratio is used in place of the likelihood ratio. As review, let us outline this procedure.

$$p(n_1, n_2, \ldots n_k \ldots n_N) = \frac{1}{\sqrt{2\pi W^2 f}}\, e^{-\frac{1}{2W^2 f} \sum_1^N n_k^2}$$

Thus

$$\frac{p(y_1, y_2, \ldots y_N / s_{01}, s_{02}, \ldots s_N)}{p(y_1 y_2, \ldots y_N / s_{11}, s_{12}, \ldots s_{1N})} = \frac{e^{-\frac{1}{2W^2 f} \sum_1^N (y_k - s_{ok})^2}}{e^{-\frac{1}{2W^2 f} \sum_1^N (y_k - s_{1k})^2}}$$

or the log of the likelihood ratio is

$$\frac{1}{2W^2\ f} \left[\sum_{k=1}^{N} (y_k - s_{ok})^2 - (y_k - s_{1k})^2 \right]$$

In the limit as $t_\Delta = \frac{1}{f} \to 0$, this becomes

$$-\frac{1}{2W^2} \left\{ \int_0^T [y(t) - s_o(t)]^2 - [y(t) - s_1(t)]^2 \right\} dt$$

in which T is the duration of the signals. This quantity is the actual voltage level "calculated" by the receiver as a basis for deciding which signal is present. If we multiply the terms out and discard the factor $\frac{1}{2} W^2$ as being a scale factor, then the decision voltage becomes

$$2 \int_0^T y(t)[s_o(t) - s_1(t)]\ dt + \int_0^T [s_1^2(t) - s_o^2(t)]\ dt$$

The significant part of this expression is

$$\int_0^T y(t)[s_o(t) - s_1(t)]\, dt$$

since this is the only component which depends on the actual signal sent. The remainder may be included in the selection of the threshold level. Now, if we let

$$f(t) = [s_o(t) - s_1(t)]$$

we obtain the result that the received signal y(t) should, for optimum detection, be correlated with f(t) and the output of the correlator (or the equivalent) be used to decide which signal was sent. For the case in which the interference is white noise, the signal to be generated at the receiver is the difference of the message signals. As we shall see for the colored noise case, the optimum receiver will be similar, but f(t) will include not only the signal structure but the noise structure as well.

II. Details of the Direct Solution

When the interference is colored noise the joint probability $p(n_1, n_2, \ldots n_k \ldots n_N)$ cannot be written as the product of the marginal probability $p(n)$ because the noise is no longer independent from one sample to another. However, we can use the fact that the noise is gaussian by using the N^{th} order joint gaussian distribution. The procedure is algebraically complex but not insurmountable. To reduce this complexity matrix notation will be used.

The joint gaussian distribution is not readily written in terms of the corresponding power spectrum, but rather in terms of the autocorrelation function. Even then the results are not obvious or easily interpreted.

Let $R(\tau)$ be the autocorrelation function of the noise. The autocorrelation function is the Fourier transform of the noise power spectrum. Let $\mathcal{R}$ be the matrix formed from the values of the autocorrelation function. That is

$$\mathcal{R} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1N} \\ R_{21} & R_{22} & \cdots & R_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ R_{N1} & R_{N2} & \cdots & R_{NN} \end{bmatrix}$$

where

$$R_{mn} = R(\tau_m - \tau_n)$$

and

τ_m corresponds to the m^{th} sample time

τ_n corresponds to the n^{th} sample time.

In addition to the above array another matrix is needed. Let

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1N} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_{N1} & \lambda_{N2} & \cdots & \lambda_{NN} \end{bmatrix}$$

in which Λ is the inverse of $\mathcal{R}$. That is if the elements of $\mathcal{R}$ were the coefficients of a set of equations such as

$$\begin{aligned} y_1 &= R_{11} X_1 + R_{12} X_2 + \dots + R_{1N} X_N \\ y_2 &= R_{21} X_1 + R_{22} X_2 + \dots + R_{2N} X_N \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_N &= R_{N1} X_1 + R_{N2} X_2 + \dots + R_{NN} X_N \end{aligned}$$

then the inverse set, or the solution, has the elements of Λ as the coefficients - i.e.

$$\begin{aligned} X_1 &= \lambda_{11} y_1 + \lambda_{12} y_2 + \dots + \lambda_{1N} y_N \\ X_2 &= \lambda_{21} y_1 + \lambda_{22} y_2 + \dots + \lambda_{2N} y_N \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ X_N &= \lambda_{N1} y_1 + \lambda_{N2} y_2 + \dots + \lambda_{NN} y_N \end{aligned}$$

Now we can write the joint gaussian distribution that the noise amplitudes at the N sample times are $n_1, n_2, \ldots n_k, \ldots n_N$ as

$$p(n_1, n_2 \ldots n_N) = \frac{1}{(2\pi)^{\frac{N}{2}} |\Lambda|^{-\frac{1}{2}}} \exp\left\{- \frac{1}{2} \sum_{m=1}^{N} \sum_{n=1}^{N} \lambda_{mn} n_m n_n\right\}$$

where

$\Lambda = \{\lambda_{mn}\}$ is the inverse of the autocorrelation matrix

$|\Lambda| =$ determinant corresponding to Λ

If, as for the case of white noise, we form the log of the likelihood ratio, the result is

$$- \frac{1}{2} \left\{\sum_{m=1}^{N} \sum_{n=1}^{N} \lambda_{mn}(y_m - s_{om})(y_n - s_{on}\right\}$$

$$+ \frac{1}{2} \left\{ \sum_{m=1}^{N} \sum_{n=1}^{N} \lambda_{mn}(y_m - s_{1m})(y_n - s_{1n})\right\}$$

The factor $\frac{1}{2}$ can be included in the decision threshold. The voltage level or the test statistic, the receiver "evaluates" is this quantity or the equivalent. By multiplying out, the test statistic can be simplified to

1

$$\sum_{m=1}^{N} \sum_{n=1}^{N} \lambda_{mn}(- y_m y_n + s_{om}\, y_n + y_m\, s_{on} - s_{om}\, s_{on}$$

$$+ y_m y_n - s_{1m}\, y_n - y_m\, s_{1n} + s_{1m}\, s_{1n})$$

Noting

$$\sum_{m=1}^{N} \sum_{n=1}^{N} s_{om}\, y_n = \sum_{m=1}^{N} \sum_{n=1}^{N} y_m\, s_{on}$$

and similarly for $s_1(t)$, the test statistic reduces to

$$\sum_{m=1}^{N} \sum_{n=1}^{N} \lambda_{mn}(2y_m\, s_{on} - 2y_m\, s_{1n} - s_{om}\, s_{on} + s_{1m}\, s_{1n})$$

The last two terms of the test statistic are deterministic. Hence the test statistic can be simplified to

$$\sum_{m=1}^{N} \sum_{n=1}^{N} \lambda_{mn}(y_m)(s_{on}-s_{1n})$$

Let

$$f_m = \sum_{n=1}^{N} \lambda_{mn}(s_{on}-s_{1n})$$

The test statistic is then

$$\sum_{m=1}^{N} f_m y_m$$

1

or as $N \to \infty$
$$\int_0^T f(t)\, y(t)\, dt$$

Thus a result is obtained for the colored noise case which is similar to that obtained for the white noise case - namely, that the optimum receiver should employ correlation. Instead of the locally generated signal at the receiver being the difference signal

$$f(t) = [s_o(t) - s_1(t)]$$

it should be determined from the sampled values,

$$f_m = \sum_{n=1}^{N} \lambda_{mn}(s_{on} - s_{1n})$$

To do this we note that the matrix $\{\lambda_{mn}\}$ and $\{R_{nm}\}$ are inverses of each other. Hence the set of simultaneous equations for the sampled values of $f(t)$ - i.e.

$$f_1 = \lambda_{11}(s_{o1} - s_{11}) + \lambda_{21}(s_{o2} - s_{12}) + \ldots \lambda_{1N}(s_{oN} - s_{1N})$$
$$f_2 = \lambda_{21}(s_{o1} - s_{11}) + \lambda_{22}(s_{o2} - s_{12}) + \ldots \lambda_{2N}(s_{oN} - s_{1N})$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$f_N = \lambda_{N1}(s_{o1} - s_{11}) + \lambda_{N2}(s_{o2} - s_{12}) + \ldots \lambda_{NN}(s_{oN} - s_{1N})$$

has the solution

$$(s_{o1} - s_{11}) = R_{11}f_1 + R_{12}f_2 + \dots R_{1N}f_N$$
$$(s_{o2} - s_{12}) = R_{21}f_1 + R_{22}f_2 + \dots R_{2N}f_N$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$(s_{oN} - s_{1N}) = R_{N1}f_1 + R_{N2}f_2 + \dots R_{NN}f_N$$

or for all n

$$(s_{on} - s_{1n}) = \sum_{m=1}^{N} R_{nm}f_m$$

In the limit as $t_\Delta \to 0$, $N \to \infty$ and

$$[s_o(t) - s_1(t)] = \int_0^T R(t-\tau)\ f(\tau)\ d\tau$$

This is the basic equation from which f(t) is determined. In the case in which the noise is white, R(t) is a delta function - i.e. $R(\tau) = 2W^2\ \delta(\tau)$. In this case

$$[s_o(t) - s_1(t)] = 2W^2\ f(t)$$

or neglecting the scale factor

$$f(t) = [s_o(t) - s_1(t)]$$

as before.

In order to design the optimum receiver it is necessary to solve the integral equation

$$[s_0(t) - s_1(t)] = \int_0^T R(t-\tau)\, f(\tau)\, d\tau$$

for the waveform f(t). Once this waveform is known the received signal can be correlated (or an equivalent operation performed) with f(t) and the output of the correlator compared with a preset threshold to decide if $s_0(t)$ or $s_1(t)$ is more likely the signal sent.

To determine f(t) we proceed in a manner analogous to solving ordinary differential equations. That is, the homogeneous equation is solved for first. The homogeneous integral equation is

$$\sigma^2\, \varphi(t) = \int_0^T R(t-\tau)\, \varphi(\tau)\, d\tau$$

This equation has the trivial solution $\varphi(t) = 0$. It also has non-trivial solutions for certain definite values of σ. Each value of σ for which there is a non-trivial solution is called an eigenvalue, and the corresponding function $\varphi(t)$ is called an eigenfunction.

Diversion

An integral equation of the form

$$\sigma^2 \varphi(t) = \int_0^T R(t-\tau)\, \varphi(\tau)\, d\tau$$

has solutions of the same character as an ordinary differential equation plus boundary conditions. For example, consider the ordinary differential equation

$$\frac{d^2\, y(t)}{dt^2} + \omega^2 y(t) = 0$$

This has the trivial solution $y(t) = 0$. It has the general non-trivial solution

$$y(t) = A \cos \omega t + B \sin \omega t$$

in which A and B are arbitrary, and ω can be any constant. If, however, we impose the boundary conditions

$$y(0) = 0$$

$$y(1) = 0$$

then the solution becomes

$$y(t) = B \sin \pi n t$$

Thus, only if the constant ω equals πn for $n = 0, \pm 1, \pm 2, \ldots$ is there a non-trivial solution to the differential equation which also satisfies the boundary conditions. These values of ω for which a non-trivial solution exists are called eigenvalues and the corresponding functions are called eigenfunctions.

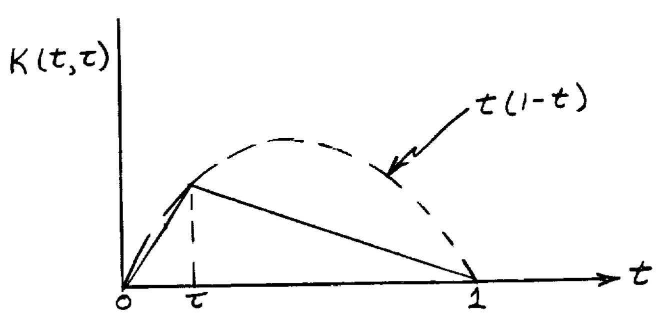
A homogeneous ordinary differential equation with boundary conditions can be converted to an integral equation. For example, consider the above differential equation with the boundary conditions. These correspond to the integral equation

$$\omega^2 \, y(t) = \int_0^T K(t,\tau) \, y(\tau) \, d\tau$$

in which $K(t,\tau)$, the kernel, is

$$\begin{aligned} K(t,\tau) &= (1-t)\,\tau & \qquad 0 \le \tau \le t \le 1 \\ &= t(1-\tau) & \qquad 0 \le t \le \tau \le 1 \end{aligned}$$

This type of kernel, called triangular, is shown below.



Because of the relation between integral and differential equations, one valuable technique for solving integral equations is to determine the corresponding differential equation.

<u>End of Diversion</u>

For convenience consider the solution of this equation corresponding to the lowest value of σ^2 to be the first eigenfunction, the next lowest, the second eigenfunction and so on. Then one interesting property of these solutions is that they are orthogonal and complete. Thus if $\varphi_k(t)$ and $\varphi_m(t)$ are two solutions

$$\int_0^T \varphi_k(t)\ \varphi_m(t) = 0 \qquad \text{if} \qquad k \neq m.$$

This property makes it easy to expand the signal waveforms in terms of these functions, for if we let

$$s_o(t) = C_1\ \varphi_1(t) + C_2\ \varphi_2(t) + C_3\ \varphi_3(t) + \ldots.$$

Then

$$\int_0^T s_o(t)\ \varphi_k(t)\ dt = C_1 \int_0^T \varphi_1(t)\ \varphi_k(t)\ dt + \ldots\ C_k \int_0^T \varphi_k^2(t)\ dt + \ldots$$

and since the φ's are orthognal

$$\int_0^T s_o(t)\ \varphi_k(t)\ dt = C_k \int_0^T \varphi_k^2(t)\ dt$$

Hence the coefficients of the expansion are given by

$$C_k = \frac{\int_0^T s_o(t)\, \varphi_k(t)\, dt}{\int_0^T \varphi_k^2(t)\, dt}$$

Note that since the equations are homogeneous the φ's can be multiplied by any factor and still remain solutions of the integral equation. By properly selecting this factor we can normalize the φ's so that

$$\int_0^T \varphi_k^2(t)\, dt = 1$$

for each function. Then

$$C_k = \int_0^T s_o(t)\, \varphi_k(t)\, dt$$

Similarly if

$$s_1(t) = d_1\, \varphi_1(t) + d_2\, \varphi_2(t) + d_3\, \varphi_3(t) + \ldots.$$

Then

$$d_k = \int_0^T s_1(t)\, \varphi_k(t)\, dt$$

If now we wish to solve the non-homogeneous integral equation for the correlation signal f(t) - i.e. solve

$$\int_0^T R(t-\tau)\ f(\tau)\ d\tau \ = \ s_o(t) - s_1(t)$$

We seek a series solution for f(t) of the same form

$$f(t) \ = \ e_1\ \varphi_1(t) + e_2\ \varphi_2(t) + \ldots .$$

We note

$$\int_0^T R(t-\tau)\ f(\tau)\ d\tau \ = \ \int_0^T R(t-\tau) \left\{ \sum_{k=1}^{\infty} e_k\ \varphi_k(\tau) \right\} \ d\tau$$

$$= \ \sum_{k=1}^{\infty} e_k \int_0^T R(t-\tau)\ \varphi_k(\tau)\ d\tau$$

But since $\varphi_k(t)$ is a solution of the homogeneous equation

$$\int_0^T R(t-\tau)\ \varphi_k(\tau) \ = \ \sigma_k^2\ \varphi_k(t)$$

$$\int_0^T R(t-\tau)\ f(\tau)\ d\tau \ = \ \sum_{k=1}^{\infty} e_k\ \sigma_k^2\ \varphi_k(t)$$

This must equal $s_o(t) - s_1(t)$, or

$$s_o(t) - s_1(t) = \sum_{k=1}^{\infty} (C_k - d_k)\, \varphi_k(t)$$

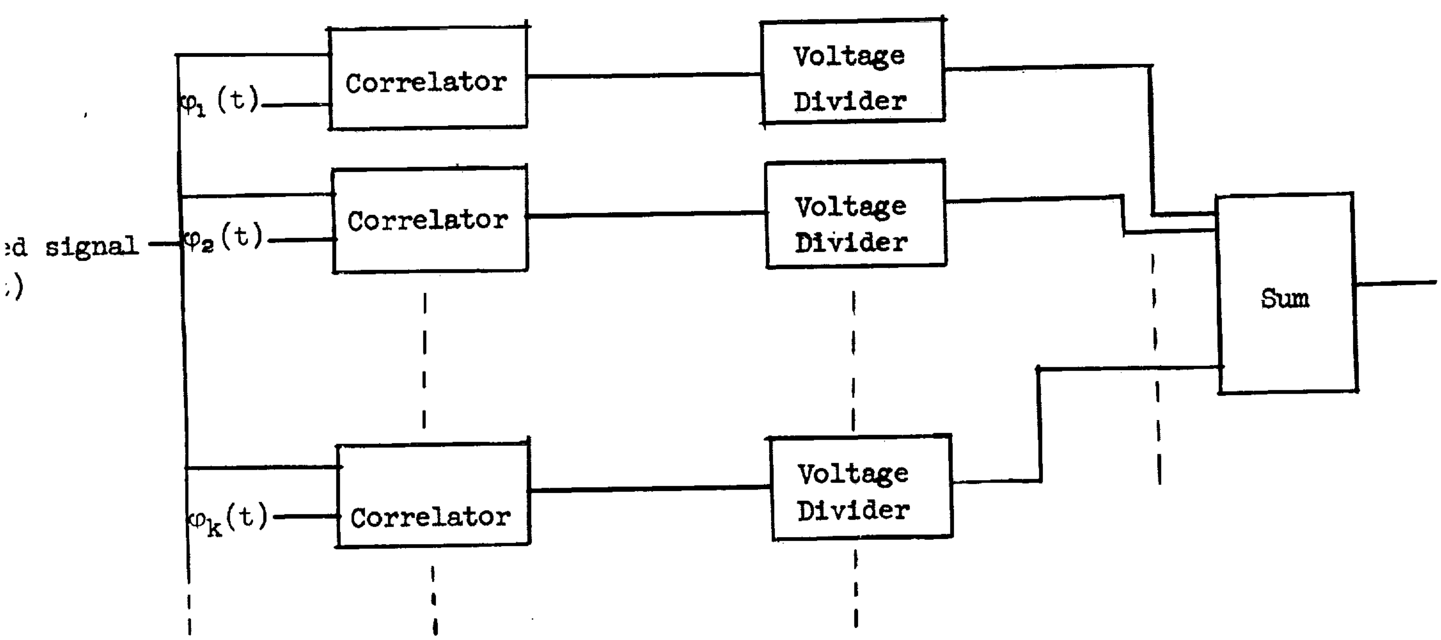
Equating coefficients we find

$$e_k = \left(\frac{C_k - d_k}{\sigma_k^2}\right)$$

or

$$f(t) = \sum_{k=1}^{\infty} \left(\frac{C_k - d_k}{\sigma_k^2}\right) \varphi_k(t)$$

From the relation one form of the optimum receiver is found to be



Each voltage divider is adjusted for the appropriate factor - the k^{th} divider for $(C_k-d_k)/\sigma_k^2$. Note that since σ_k^2 increases for increasing k an infinite number of correlators and dividers is not required for a practical system as the contribution to the sum becomes very small for high k. It can be shown that σ_k^2 is the variance or mean squared noise of the output of the k^{th} correlator. Thus the high noise correlator outputs are properly weighted down by the divider circuits.

III. Details of the Indirect Solution

In order to gain further insight into the basis and operation of an optimum receiver for detecting signals in colored noise, let us start anew and consider this problem independently of the previous analysis.

Consider that one of two known signals $s_0(t)$ and $s_1(t)$ is transmitted for a fixed interval $0 \leq t < T$. The transmitted signal is corrupted by additive stationary gaussian noise of a known spectral density or, correspondingly, a known autocorrelation function. The decision to be made at the receiver is whether $s_0(t)$ or $s_1(t)$ was actually sent.

We could proceed as before and base the likelihood ratio on the values attained by the received signal $y(t)$ at intervals of time spaced t_Δ apart ($t_\Delta \to 0$ in the limit). However, these observed values being correlated led to the algebraic difficulties previously considered. Specifically, the joint conditional probability $p(y_1, y_2, y_3 \ldots / s_{01}, s_{02}, s_{03} \ldots)$ cannot be written as the product of probabilities of the form $p(y_j/s_{0j})$ unless the noise is white.

Now the set of values $(y_1, y_2 \ldots)$ can be looked upon as a set of observable coefficients or coordinates on which the likelihood ratio is based. Note that the set to use is <u>not</u> suggested by the likelihood ratio and must be independently sought.

The set of coefficients or coordinates previously used was the sampled values of the received signal (in the limit the instantaneous signal amplitude). Here we shall use what appears to be a different set of observable coordinates and later show they are in effect the same.

One useful suggestion is to find a set of observable coordinates y_k that are uncorrelated but can be generated from the received signal $y(t)$ by linear operations. We desire to have

$$y(t) = \sum_k y_k \Phi_k(t)$$

in which, for convenience, the set $\Phi_k(t)$ is orthonormal with respect to the internal $0 \leq t \leq T$. This permits the coordinates (or coefficients) y_k to be computed from

$$y_k = \int_0^T \Phi_k(t)\, y(t)\, dt$$

Thus the optimum receiver will be as previously shown, but the voltage division ratios remain to be determined. In the receiver any one of the coefficients y_k could be generated by passing the received signal $y(t)$ through a filter matched to the waveform $\Phi_k(t)$, $0 \leq t < T$. At the end of the observation interval the output of the filter would be the coefficient y_k for the received signal $y(t)$. Note that y_k is a random variable in the sense that if $s_o(t)$ is repeatedly sent, then the different waveforms the random noise interference takes yields different values for y_k. Thus y_k has a certain probability distribution when $s_o(t)$ is sent and a different (we hope) probability distribution when $s_2(t)$ is sent.

One possible orthonormal set is $\left\{ \sqrt{\frac{1}{T}}, \sqrt{\frac{2}{T}} \cos \frac{2\pi t}{T}, \sqrt{\frac{2}{T}} \sin \frac{2\pi t}{T}, \ldots \sqrt{\frac{2}{T}} \cos \frac{2\pi n t}{T}, \sqrt{\frac{2}{T}} \sin \frac{2\pi n t}{T}, \ldots \right\}$

However we not only desire the convenience of an orthonormal set, but that the set $\{y_k\}$ corresponding to $\left\{ \begin{matrix} s_1(t) \\ s_o(t) \end{matrix} \right\}$ being sent be composed of random variables which are independent. Then,

$$p[y_1, y_2, \ldots\ldots / s_o(t)] = p[y_1/s_o(t)]\, p[y_2/s_o(t)] \ldots\ldots$$

which is the relation desired.

We note, since the noise is additive, that

$$y(t) = s(t) + n(t)$$

in which $s(t)$ stands for either $s_o(t)$ or $s_1(t)$, and $n(t)$ is the additive gaussian interference. Then

$$y_k = \int_0^T \Phi_k(t)\, y(t)\, dt$$

$$y_k = \int_0^T \Phi_k(t)\, s(t)\, dt + \int_0^T \Phi_k(t)\, n(t)\, dt$$

Thus y_k has a gaussian distribution. The mean value is

$$\text{mean } y_k = \int_0^T \Phi_k(t)\ s(t)\ dt$$

and the variance is

$$\text{variance } y_k = < \left[\int_0^T \Phi_k(t)\ n(t)\ dt \right]^2 >$$

$$= \int_0^T \int_0^T \Phi_k(t)\ \Phi_k(\tau) < n(t)\ n(\tau) > dt\ d\tau$$

$$= \int_0^T \int_0^T \Phi_k(t)\ \Phi_k(\tau)\ R(t-\tau)\ dt\ d\tau$$

in which $R(t-\tau)$ is the autocorrelation function of the noise. Now we desire y_j, y_k to be independent for $j \neq k$ in the sense that the random parts of y_j, y_k are to be independent. This means

$$< \int_0^T \Phi_k(t)\ n(t)\ dt \cdot \int_0^T \Phi_j(\tau)\ n(\tau)\ d\tau > = 0$$

or

$$\int_0^T \int_0^T \Phi_k(t)\ \Phi_j(\tau) < n(t)\ n(\tau) > dt\ d\tau = 0$$

$$\text{or} \qquad \int_0^T \int_0^T \Phi_k(t)\ \Phi_j(t)\ R(t-\tau)\ dt\ d\tau = 0 \qquad \text{for } j \neq k$$

Assume that we can solve the above equation for an orthonormal set $\{\Phi_k(t)\}$ such that the observable coordinates are independent. Further assume that the set $\{\Phi_k(t)\}$ is also complete. Then the receiver can generate, by means of a bank of filters matched to the set $\{\Phi_k(t)\}$, a set of outputs y_k at the end of the observation interval. As can be seen the output y_k can be written as

$$y_k = a_k + z_k \quad \text{if} \quad s_0 \text{ is sent}$$

or

$$y_k = b_k + z_k \quad \text{if} \quad s_1 \text{ is sent}$$

in which

$$a_k = \int_0^T s_0(t)\ \Phi_k(t)\ dt$$

$$b_k = \int_0^T s_1(t)\ \Phi_k(t)\ dt$$

and z_k is a random variable of zero mean, different z_k's being uncorrelated and hence the z_k's are mutually independent gaussian random variables. As before the likelihood ratio can be taken and again, for convenience, the natural logarithm of the likelihood ratio considered as the test statistic. This is, using the first

N coordinates,

$$\tfrac{1}{2}\left\{ - \sum_{k=1}^{N} \frac{(a_k - y_k)^2}{\sigma_k^2} + \sum_{k=1}^{N} \frac{(b_k - y_k)^2}{\sigma_k^2} \right\}$$

in which σ_k^2 is the variance of z_k.

The analogy to the white gaussian noise problem is now quite evident. The main problem is to generate the orthonormal set $\{\Phi_k(t)\}$. To do this we note that the equation

$$\int_0^T \int_0^T \Phi_k(t)\ \Phi_j(\tau)\ R(t-\tau)\ dt\ d\tau = 0\ , \quad j \neq k$$

holds true if $\displaystyle\int_0^T \Phi_j(\tau)\ R(t-\tau)\ d\tau = \lambda_j\ \Phi_j(t)$

for then

$$\int_0^T \int_0^T \Phi_k(t)\ \Phi_j(\tau)\ R(t-\tau)\ dt\ d\tau = \int_0^T \Phi_k(t) \cdot \lambda_j\ \Phi_j(t)\ dt$$

$$= 0 \quad \text{if } k \neq j$$

$$= \lambda_j \quad \text{if } k = j$$

Since $\displaystyle\int_0^T \int_0^T \Phi_j(t)\ \Phi_j(\tau)\ R(t-\tau)\ dt\ d\tau = \sigma_j^2$, the variance of z_j,

then $\lambda_j = \sigma_j^2$

Hence to find the orthonormal functions $\Phi_j(t)$ we need to solve the homogeneous equation

$$\int_0^T \Phi_j(\tau)\ R(t-\tau)\ d\tau = \sigma_j^2\ \Phi_j(t)$$

in which $R(\tau)$ is the autocorrelation function of the additive gaussian noise interference.

To find the actual quantity the receiver is to evaluate, note that by using the first N coordinates the log of the likelihood ratio is

$$\tfrac{1}{2}\left\{ -\sum_{k=1}^{N} \frac{(a_k - y_k)^2}{\sigma_k^2} + \sum_{k=1}^{N} \frac{(b_k - y_k)^2}{\sigma_k^2} \right\}$$

or
$$\tfrac{1}{2}\sum_{k=1}^{N} \frac{b_k^2 - a_k^2}{\sigma_k^2} + \sum_{k=1}^{N} \frac{y_k(a_k - b_k)}{\sigma_k^2}$$

The first term is a deterministic, since this depends only on the signal structure - i.e.,

$$a_k = \int_0^T s_0(t)\ \Phi_k(t)\ dt$$

$$b_k = \int_0^T s_1(t)\ \Phi_k(t)\ dt$$

The value of the second term depends on which signal is sent and, of course, the noise. Therefore the likelihood ratio test is based on comparing

$$\sum_{k=1}^{N} \frac{y_k\,(a_k - b_k)}{\sigma_k^2}$$

with a threshold.

Before considering the limiting behavior as $N \to \infty$, it is convenient to introduce

$$f_N(t) = \sum_{k=1}^{N} \frac{(a_k - b_k)}{\sigma_k^2}\,\Phi_k(t)$$

If σ_k^2 was independent of k, for example $\sigma_k^2 = 1$, then

$$\lim_{N \to \infty} f_N(t) = \sum_{k=1}^{\infty} a_k\,\Phi_k(t) - \sum_{k=1}^{\infty} b_k\,\Phi_k(t)$$

$$= s_0(t) - s_1(t)$$

Otherwise the relation is not as simple.

To write the test statistic in terms of $f_N(t)$ we note that

$$\frac{1}{2} \sum_{k=1}^{N} \frac{(b_k^2 - a_k^2)}{\sigma_k^2} + \sum_{k=1}^{N} \frac{y_k(a_k - b_k)}{\sigma_k^2}$$

can be written as

$$\tfrac{1}{2}\sum_{k=1}^{N} \frac{(b_k + a_k)(b_k - a_k)}{\sigma_k^2} + \sum_{k=1}^{N} \frac{y_k(a_k - b_k)}{\sigma_k^2}$$

or $$\tfrac{1}{2}\sum_{k=1}^{N} \frac{\left\{ \int_0^T [s_1(t) + s_o(t)]\, \Phi_k(t)\, dt \right\}\left\{ b_k - a_k \right\}}{\sigma_k^2}$$

$$+ \sum_{k=1}^{N} \frac{\left[\int_0^T y(t)\, \Phi_k(t)\, dt\right]\left[a_k - b_k\right]}{\sigma_k^2}$$

or $$\int_0^T \left[y(t) - \frac{s_o(t) + s_1(t)}{2}\right] \underbrace{\sum_{k=1}^{N} \Phi_k(t) \frac{[a_k - b_k]}{\sigma_k^2}}_{\text{This is } f_N(t)} dt$$

or $$\int_0^T f_N(t) \left[y(t) - \frac{s_o(t) + s_1(t)}{2}\right] dt$$

Formally if $N \to \infty$, the test statistic becomes

$$\int_0^T f(t) \left[y(t) - \frac{s_o(t) + s_1(t)}{2}\right] dt$$

Note that since $f(t)$, $s_o(t)$, and $s_1(t)$ are deterministic, the test statistic can be simplified to

$$\int_0^T y(t)\ f(t)\ dt$$

Thus detection in colored noise is by correlation. Instead of correlating the received signal with $[s_0(t) - s_1(t)]$ as for white gaussian noise, however, the related signal $f(t)$ is used. To determine the integral equation for $f(t)$, we note, by definition,

$$f_N(t) = \sum_{k=1}^{N} \frac{(a_k - b_k)}{\sigma_k^2}\ \Phi_k(t)$$

and that $f_N(t)$ is thus the solution of

$$\int_0^T R(t-\tau)\ f_N(\tau)\ d\tau = \sum_{k=1}^{N} \int_0^T R(t-\tau)\ \frac{(a_k - b_k)}{\sigma_k^2}\ \Phi_k(\tau)\ d\tau$$

$$= \sum_{k=1}^{N} \frac{(a_k - b_k)}{\sigma_k^2} \underbrace{\int_0^T R(t-\tau)\ \Phi_k(\tau)\ d\tau}_{\substack{\text{by definition of } \Phi_k(t) \\ \equiv \sigma_k^2\ \Phi_k(t)}}$$

$$= \sum_{k=1}^{N} (a_k - b_k)\ \Phi_k(t)$$

As $N \to \infty$, at least formally, $f(t)$ is the solution of:

$$\int_0^T R(t-\tau)\, f(\tau)\, d\tau = s_0(t) - s_1(t)$$

This result, together with the fact that the receiver is to compute

$$\int_0^T y(t)\, f(t)\, dt$$

specifies the optimum receiver.

3

IV. A Simple Example

Consider the case in which the power spectral density of the noise is

$$W_n(\omega) = A^2 + B^2 \omega^2 \quad \text{(double-sided spectrum)}$$

This can be considered the "video-noise" corresponding to the synchronous detection of an r-f signal, or a "weighting function" representing the receiver bandpass characteristics.

Then
$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} [A^2 + B^2 \omega^2] \, e^{j\omega\tau} \, d\omega$$

$$= A^2 \, \delta(\tau) - B^2 \, \delta''(\tau)$$

in which $\delta(\tau)$ = Dirac delta function

$\delta''(\tau)$ = Second derivative of the Dirac delta function.

The orthonormal functions $\Phi_k(t)$ are solutions of

$$\sigma_k^2 \, \Phi_k(t) = \int_0^T \Phi_k(\tau) \, R(t-\tau) \, d\tau$$

or $$\sigma_k^2 \, \Phi_k(t) = \int_0^T \Phi_k(\tau) \, [A^2 \, \delta(t-\tau) - B^2 \, \delta''(t-\tau)] \, d\tau$$

$$= A^2 \, \Phi_k(t) - B^2 \, \Phi_k''(t)$$

Thus the functions $\Phi_k(t)$ are solutions of the differential equation

$$B^2 \, \Phi_k''(t) + (\sigma_k^2 - A^2) \, \Phi_k(t) = 0$$

subject to the condition

$$\int_0^T \Phi_k(t) \, \Phi_n(t) \, dt = \delta_{kn}$$

The solutions are

$$k_1 \cos t \sqrt{\frac{\sigma_k^2 - A^2}{B^2}} \text{ and } k_2 \sin t \sqrt{\frac{\sigma_k^2 - A^2}{B^2}}$$

the requirement

$$\int_0^T \Phi_k \, \Phi_n \, dt = \delta_{kn}$$

giving

$$\sqrt{\frac{\sigma_k^2 - A^2}{B^2}} = \frac{2\pi k}{T}$$

or $$\sigma_k^2 = A^2 + \left(\frac{2\pi k}{T} \right) B^2$$

The functions $\Phi_k(t)$ are thus of the normalized form

$$\sqrt{2}\cos\frac{2\pi kt}{T} \quad \text{and} \quad \sqrt{2}\sin\frac{2\pi kt}{T}$$

The k^{th} observable coordinates are therefore

$$\sqrt{2}\int_0^T y(t)\cos\frac{2\pi kt}{T}\,dt \quad \text{and} \quad \sqrt{2}\int_0^T y(t)\sin\frac{2\pi kt}{T}\,dt$$

These observable coordinates are proportional to the Fourier Series coefficients of the received signal. Note, however, that the variance associated with the k^{th} coordinates, σ_k^2, increases with increasing k. This is reflected in the fact that higher order coordinates are more heavily weighted before being summed to form the output level of the receiver (See Figure).